



Original communication

The Weighted Average Method ‘WAM’ for Dental Age Estimation: A simpler method for children at the 10 year threshold



“It is vain to do with more when less will suffice” William of Ockham 1288–1358.”[☆]

Graham J. Roberts^{a,*}, Fraser McDonald^a, Monica Neil^{a,b}, Victoria S. Lucas^a

^a Department of Orthodontics, King's College London Dental Institute, Floor 22 Tower Wing, Guy's Hospital, London SE1 9RT, UK

^b Department of Maxillo-Facial and Dental Surgery, Great Ormond Street Hospital, Great Ormond Street, London WC1N 3JH, UK

ARTICLE INFO

Article history:

Received 25 February 2014

Accepted 16 May 2014

Available online 10 June 2014

Keywords:

Dental Age Estimation
Dental Age Assessment
Weighted average
Meta-analysis
Method comparison

ABSTRACT

The mathematical principle of weighting averages to determine the most appropriate numerical outcome is well established in economic and social studies. It has seen little application in forensic dentistry.

This study re-evaluated the data from a previous study of age assessment at the 10 year threshold.

A semiautomatic process of weighting averages by n-tds, x-tds, sd-tds, se-tds, 1/sd-tds, 1/se-tds was prepared in an Excel worksheet and the different weighted mean values reported. In addition the Fixed Effects and Random Effects models for Meta-Analysis were used and applied to the same data sets.

In conclusion it has been shown that the most accurate age estimation method is to use the Random Effects Model for the mathematical procedures.

© 2014 Elsevier Ltd and Faculty of Forensic and Legal Medicine. All rights reserved.

1. Introduction

Age assessment using radiographically discernible stages of tooth development has been a practical method of age estimation since the early 1960's.¹ In essence, the various methods used consist of the development of a Reference Data Set which provides summary data for each Tooth Development Stage (TDS). This consists of the count for a given tooth development stage, the mean, and the standard deviation conveniently identified as n-tds, x-tds, and sd-tds. The data for each of the TDSs are used in a variety of statistical techniques to carry out an age estimation. These vary from a simple median,² through weighted scores,³ cluster analysis,⁴ multivariate analysis,⁴ linear regression,⁵ multiple regression,⁶ logistic regression,⁷ and Bayesian inference.⁸ The clinician carrying out Dental Age Assessments (DAA), usually *not* a statistician, and

attempting to comprehend these different approaches, becomes confused at the prospect of identifying which is the most accurate method.

A recent study reported that, on average, the DA was within 3 months of the chronological age.⁹ This study was described as a ‘Simple Method’. The calculation for DA was based on the weighted average obtained using the mathematical methods of meta-analysis.⁹ Much time and discussion was expended to determine whether the Fixed Effects or Random Effects calculation should be used¹⁰ and whether the weighting factor should be the standard deviation (sd-tds) or the standard error (se-tds) of the Age at Attainment (AaA) for each Tooth Development Stage (TDS). This strategy was adopted and used for a number of investigations reporting a comparison between Chronological Age (CA) and Dental Age (DA) at the 10 year threshold,¹¹ 13 year threshold,¹² 16 year threshold,¹³ and a general age range.⁹ These gave good results with the average of the CA *minus* DA difference being 3 months or less. This was reported as the most reliable estimate of DA for validation studies on children, adolescents and emerging adults. An underlying difficulty with this simple method was the fact that special statistical software is needed to calculate the DA.^{14,15} As a consequence, clinicians carrying out Age Assessments were reluctant to use the method.

[☆] This principle is variously formulated. In essence it is a scientific and philosophical rule that entities should not be multiplied unnecessarily which is interpreted as requiring that the simplest of competing theories be preferred to the more complex or that explanations of unknown phenomena be sought first in terms of known quantities – Wikipedia 2012.

* Corresponding author. Tel.: +44 020 7188 4432; fax: +44 020 7188 4415.
E-mail address: graham.j.roberts@kcl.ac.uk (G.J. Roberts).



Fig. 1. Dental Panoramic Tomograph – Female subject of known age (study group).

Further consideration led to the conjecture that exploration of the concept of the Weighted Average Method would provide a full understanding of the effects of different weighting factors. A compelling aspect of this is that widely used software has a number of mathematical and statistical functions that make the calculation of summary statistical data and weighted averages a matter of simple routine.¹⁶

This paper assesses the effect of different weighting factors on the accuracy of Dental Age Estimates using simple weightings and comparing them with meta-analysis outcomes.

This is, in effect, a series of validation tests to determine the most accurate method for estimating Dental Age using the Gold Standard of Chronological Age as the comparator.

2. Material and methods

The quantitative data on Tooth Development Stages for this study are re-used from a paper published in 2011.¹¹

To calculate the weighted average of the Ages at Attainment of the TDSs for an individual study subject, the following approach was used.

2.1. Procedure

Age estimation followed the same procedure as described in 2008.^{9,17} In brief, a small table is created using the TDS for each tooth present on the study subject's Dental Panoramic Tomograph (Fig. 1).

Each of the teeth present was graded using the 8 stage system of Demirjian.³

The n-tds, x-tds, and sd-tds for each TDS present were extracted from the Excel spreadsheet comprising the Reference Data Set¹⁷ (see Appendix 1). The TDSs and the associated data for a single case are shown in Table 1.

Table 1 also shows the Tooth Development Stages for a single individual with the associated summary data derived from the RDS entered into the columns on the right (n-tds, x-tds, and sd-tds). The Microsoft Excel spreadsheet is configured to automatically calculate the average age with the different weightings indicated in Table 1. These are in the lower half of the table.

To calculate the weighted average the following formula was used:

$$\text{Weighted Average} = ((X_1 * W_1) + (X_2 * W_2 \dots)) \div (W_1 + W_2 \dots) \text{ [18, p22]}$$

This was applied to each of the weighted averages in [c] to [g] above.

Table 1

Table created with data from the Reference Data Set (n-tds, x-tds, sd-tds) for each of the Tooth Developments Stages discernible in the Dental Panoramic Tomograph for the female subject of unknown age (SUA). The calculations are carried out automatically in the Microsoft Excel spreadsheet. The lower half of the table shows these calculated ages (a to k).

Tooth (British Dental Journal Notation & FDI Notation)	Stage	n-tds (count)	x-tds (yrs)	sd-tds (yrs)
UL1 (21)	—	—	—	—
UL2 (22)	G	18	9.06	1.34
UL3 (23)	F	28	8.53	0.85
UL4 (24)	E	19	8.08	1.35
UL5 (25)	E	20	8.61	2.12
UL6 (26)	—	—	—	—
UL7 (27)	D	50	7.87	1.57
UL8 (28)	—	—	—	—
LL8 (38)	—	—	—	—
LL7 (37)	E	13	9.06	0.93
LL6 (36)	—	—	—	—
LL5 (35)	E	25	8.76	2.33
LL4 (34)	F	27	10.27	2.64
LL3 (33)	F	32	8.36	0.92
LL2 (32)	—	—	—	—
LL1 (31)	—	—	—	—
Calculated estimates of age		Male	Years	
Chronological Age			9.22	[a]
No Weighting		[nil]	8.73	[b]
Weighted Average		[n-tds]	8.63	[c]
Weighted Average		[sd-tds]	8.04	[d]
Weighted Average		[1/sd-tds]	7.85	[e]
Weighted Average		[se-tds]	8.88	[f]
Weighted Average		[1/se-tds]	8.68	[g]
Meta-analysis Average		[sd-fixed]	8.93	[h]
Meta-analysis Average		[sd-random]	8.93	[i]
Meta-analysis Average		[se-fixed]	8.82	[j]
Meta-analysis Average		[se-random]	9.23	[k]

Table 2

Descriptions of weighting factors for the 10 comparators to chronological age.

a. Chronological age.	The Gold Standard for this study
b. No weighting.	The Average of all of the x-tds present in the study subject
c. Weighted average by n-tds.	The average determined using n-tds as the weighting factor.
d. Weighted average by sd-tds.	The average determined using sd-tds as the weighting factor.
e. Weighted Average using the reciprocal of sd-tds (1/sd-tds) as the weighting factor	
f. Weighted average by se-tds.	The average determined using se-tds as the weighting factor.
g. Weighted average using the reciprocal of se-tds (1/se-tds) as the weighting factor	
h. Meta-analysis calculation using the se-tds as the weighting factor and the fixed effects calculation routine.	
i. Meta-analysis calculation using the se-tds as the weighting factor and the random effects calculation routine.	
j. Meta-analysis calculation using the sd-tds as the weighting factor and the random effects calculation routine.	
k. Meta-analysis calculation using the se-tds as the weighting factor and the random effects calculation routine.	

For the Meta-analysis, both the fixed effects routine and the random effects routine were used with se-tds and sd-tds respectively as the weighting factor.¹⁴ The total number of ‘average ages’ reported for each subject comprised 10 estimates [for comparison with the Chronological Age. (Table 1)] (see Table 2).

A further spread sheet was created with 11 columns of data comprising the age estimates for the 50 female subjects and male subjects separately. The eleven columns contained the age estimates for a to k above. The spreadsheet had been constructed in such a way that the calculations for chronological age and weighted ages were carried out automatically. The meta-analysis calculations were carried out by copying the data from Table 1 above and pasting the table into STATA for h to k above. The calculated ages were then copied back into the Excel spreadsheet.

The ten different age estimates were then subjected to method comparison tests using the approach of Bland and Altman.¹⁹ This comprises a *t*-test to estimate the significance of the bias when the DA is subtracted from the CA. A plot of the average of the data for the two groups together (*x* axis) against the differences from the combined average (*y* axis). This shows the direction of the difference and whether there is any trend in the differences.

3. Results

The summary data for all the ten age estimation procedures are shown in Table 3. The differences in the average values between each of the weighted age estimates and the chronological age, tested using student's *t* test are shown in Table 3.

From the above data a further Table has been created which for each Assessment Method shows the mean difference between the CA and each of the weighting methods. These differences are shown in Table 4.

From Table 4 it can be seen that the most accurate results for females for simple weighting come from the comparison of ca v 1/sd [†] which is on average an overestimate of age of –0.13 years [6.76 weeks]. The comparison of ca with meta-analysis of ca versus the random effects model with se as the weighting factor [††] gives an average under estimate of 0.03 years [1.56 weeks].

The Ladder plot (Fig. 2) shows the CA (Left) and Unweighted Average (Right). This is emblematic of all the ladder plot comparisons in this study with the estimated ages (Right) showing a greater spread of years than the CA (Left). The Bland–Altman plot shows an even distribution of data above and below the zero difference line with a slight tendency to underestimate in the lower ages with slightly increased dispersion in the upper ages (Fig. 3).

Fig. 2 shows the CA on the left and the DA on the right. The feature of this is that the DA values are spread out further along the age range than the CA values. That is the DA values are less precise than the CA values. It is clear that some of the DA values are larger than the CA values and vice versa, the CA values are larger than the DA values. By inspecting this plot the impression is given that this is evenly distributed so any given radiograph from a single child will as likely result in an overestimate as an underestimate.

A detailed assessment of these differences was carried out for each result using Bland–Altman plots (Fig. 2).

Overall, below ten years of age, the preponderance of values for the difference the between the arrays of data show a slight tendency to overestimate the age. Above 10 years of age there is a tendency for underestimation of age.

4. Discussion

This study was carried out as the result if a query raised by a colleague who asked ‘what is wrong with a simple average?’

Table 3
Summary data for the 10 weighted age estimates.

Age assessment method – females	N	Mean (X̄) (years)	Standard deviation (SD) (years)	Standard error (SE) (years)
a chronological age	50	10.14	0.73	0.10
<i>Differences from CA</i>				
b Unweighted average	50	10.44	1.13	0.16
c Weighted average by n-tds	50	10.79	1.62	0.23
d Weighted average by sd-tds	50	10.60	1.14	0.16
e Weighted average by 1/sd-tds	50	10.27	1.19	0.17
f Weighted average by se-tds	50	10.47	1.01	0.14
g Weighted average by 1/se-tds	50	10.44	1.41	0.19
h Meta-analysis weighted by se-tds fixed effects	50	10.50	1.70	0.24
i Meta-analysis weighted by se-tds random effects	50	10.44	1.19	0.17
j Meta-analysis weighted by sd-tds fixed effects	50	10.18	1.26	0.18
k Meta-analysis weighted by sd-tds random effects	50	10.11	1.22	0.17
Age assessment method –males	N	Mean (X̄)	Standard Deviation (SD)	Standard Error (SE)
a Chronological age	50	9.78	0.78	0.11
<i>Differences from CA</i>				
b Unweighted average	50	9.99	1.27	0.18
c Weighted average by n-tds	50	10.04	1.45	0.20
d Weighted average by sd-tds	50	10.08	1.07	0.15
e Weighted average by 1/sd-tds	50	9.68	1.23	0.17
f Weighted average by se-tds	50	10.23	0.87	0.14
g Weighted average by 1/se-tds	50	9.75	1.35	0.19
h Meta-analysis weighted by se-tds fixed effects	50	9.63	1.44	0.20
i Meta-analysis weighted by se-tds random effects	50	9.82	1.11	0.16
j Meta-analysis weighted by sd-tds fixed effects	50	9.54	1.21	0.17
k Meta-analysis weighted by sd-tds random effects	50	9.57	1.23	0.17

[Marsden PH 2012]. The intention was to compare only an un-weighted average with the Random Effects Meta-analysis weighted by the standard error as has been reported.^{9,11–13} Detailed discussion showed that the choice of weighting factors previously used was the standard error, which incorporated both the standard deviation (sd-tds) and the number of TDSs in the Reference Data Sample (n-tds). The weighted average has an application where the investigator wishes to arrive at “An average of quantities to which have been attached a series of weights in order to make proper allowance for their relative importance. For example, a set of mean values may be combined using a weighted average in which the weights are the sample sizes on which each mean is based.”²⁰ It is clear from the data set used in this study that weighting by n-tds (“... the sample sizes ...”) gave the poorest estimate of the CA. This raises the question of how, on theoretical grounds, a weighting factor is selected and justified.

Previously we had used the standard error and justified this because it incorporates the sample size (n-tds) and the variation in the data (sd-tds). This thought process was taken further by utilising the mathematical procedures within the process of Meta-analysis.¹⁰ At the time of planning the first of our studies, it was believed that the weighted average calculated by using the se-tds and a random effects procedure would give the ‘best’ result.²⁰

At this late juncture, it is difficult to understand why we did this! What is clear is that an objective assessment of the weighted calculations has shown that the most accurate method of calculating the DA for a subject is not as clear cut as at first anticipated. The use of 50 female subjects and 50 male subjects with the Gold Standard of known CA has led to surprising results when testing the validity of the weighting factors available to investigators.

Table 4

Results of comparison between Chronological Age and Dental Ages Estimated in females and males for the 10 different methods (b to k from Table 3).

Comparison Females		N	Difference in years (bias) CA minus DA	p Value	Significance
a v b	[ca v uwa]	50	−0.30	0.0177	*
a v c	[ca v n-tds]	50	−0.65	0.0004	***
a v d	[ca v sd-tds]	50	−0.46	0.0002	***
a v e	[ca v 1/sd-tds]†	50	−0.13	0.2650	ns
a v f	[ca v se-tds]	50	−0.33	0.0023	**
a v g	[ca v 1/se-tds]	50	−0.30	0.0424	*
a v h	[ca v meta-se-fix]	50	−0.36	0.0493	*
a v i	[ca v meta-se-rnd]	50	−0.39	0.0114	*
a v j	[ca v meta-sd-fix]‡	50	0.03	0.8147	ns
a v k	[ca v meta-sd-rnd]	50	−0.04	0.7640	ns

Comparison Males		N	Difference in Years (bias) CA minus DA	p Value	Significance
a v b	[ca v uwa]‡	50	−0.12	0.3506	ns
a v c	[ca v n-tds]	50	−0.16	0.2828	ns
a v d	[ca v sd-tds]	50	−0.21	0.0365	*
a v e	[ca v 1/sd-tds]	50	0.19	0.1050	ns
a v f	[ca v se-tds]	50	−0.36	0.0009	***
a v g	[ca v 1/se-tds]	50	0.12	0.3772	ns
a v h	[ca v meta-se-fix]	50	0.24	0.1064	ns
a v i	[ca v meta-se-rnd]‡‡	50	0.06	0.6002	ns
a v j	[ca v meta-sd-fix]	50	0.31	0.0193	*
a v k	[ca v meta-sd-rnd]	50	0.34	0.0044	***

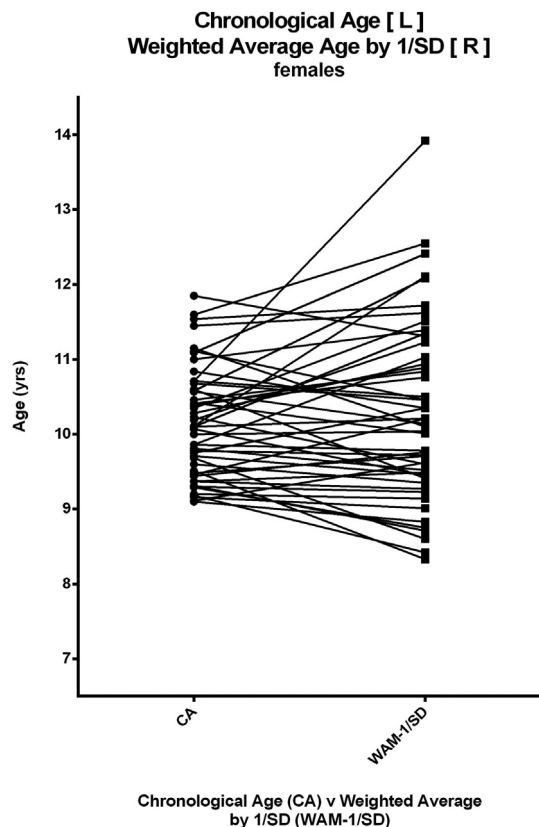


Fig. 2. This shows the CA on the left and the DA on the right. The feature of this is that the DA values are spread out further along the age range than the CA values. That is the DA values are less precise than the CA values. It is clear that some of the DA values are larger than the CA values and vice versa the CA values are larger than the DA values. By inspecting this plot the impression is given that this is evenly distributed so any given radiograph from a single child will as likely result in an overestimate as an underestimate.

Bland-Altman Plot of Chronological Age v Weighted Average Age by 1/SD

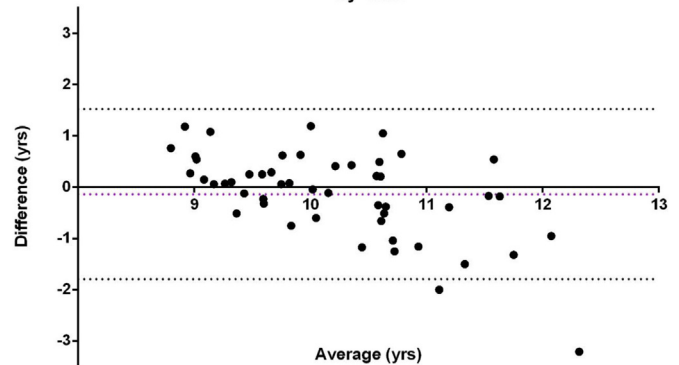


Fig. 3. Bland Altman plot of CA versus 'Average Weighted by 1/SD' in females. The black dotted lines indicate the Limits within which lie 95% of the differences between CA minus DA. The purple dotted line shows the mean difference which is extremely low (−0.13 years = 6.76 weeks).

The overall best approach is to use the meta-analysis software which gives an accuracy of 0.03 years (approximately 1.56 weeks under estimate on average) for females using the sd as the weighting factor. For males, it is the se that returns the most accurate result which is 0.06 or 3.12 weeks. These are surprisingly good results. So much so that all the radiographs and the statistical procedures have been double checked to ensure consistency.

A difficulty with using meta-analysis is that the software is expensive and many clinicians, untrained in its use and reluctant to commit to the expense continue to favour the simple un-weighted average. There is merit in this approach as the software available as part of the Microsoft Office Suite [Excel] provides a simple and effective way of semi-automating the calculations. On this basis, for females, the reciprocal of the sd returns a good result, an overestimate of −0.13 years or −6.76 weeks overestimate. For males the unweighted average also gives a good result with an overestimate of −0.12 years or 6.24 weeks.

It is not surprising that other work in this field gave less good results with the CA minus DA difference varying from 0.25 yrs to −1.19 yrs.²¹ These differences are less satisfactory than the data presented in this paper. This may be because a), data from UK Caucasian children and data from UK Bangladeshi children were merged into one data set, and b) the age estimates were carried out on the same subjects as the Reference Data Set. These approaches on logical grounds alone will lead to a less than desirable outcome. The approach used in the present study is rigorous in that only a single Identifiable Human Group has been used to create the Reference Data Set. To test the validity of the RDS it is necessary to acquire a separate study or validation group of subjects of known age. These are assessed when 'masked' and the DA compared to the CA as has been the approach in this paper. A small note of procedure is that the DA is *always* subtracted from the CA. The consequence of this is that a positive number always indicates an underestimate of the age. Conversely, a negative difference indicates an overestimate of age. A brief survey of the literature shows that several authors use the DA – CA approach. This is confusing to readers and in this paper all CA/DA differences are calculated by subtracting DA from CA [CA minus DA].

This work shows the large variation in age estimation that results from different weighting approaches.

This is compelling and objective evidence of the need to test in a systematic way any mathematical or statistical procedure by which age estimates are made. It is, perhaps, the almost unique availability of large numbers of Dental Panoramic Radiographs taken for clinical purposes that are available for re-use that enables this process of

validation to be carried out to a high level of reliability using the different methods of estimating age. Because the RDS is created from archived radiographs it is easily achievable to derive large Reference Data Sets and then test the validity of the RDS by creating suitably large Validation Samples. It is current practice to set the size of these study samples at 50 females and 50 males.

5. Conclusions

The Weighted Average Method using the mathematical procedures of meta-analysis give the most accurate results for age estimation in both females and males using the se and sd respectively as weighting factors.

For clinicians without the meta-analysis facility the use of simple weighting procedures in a semi-automated spread sheet (Excel) provides results within 4 weeks of the best results for both females and males.

There is a need for re-analysis of our own data^{12,13} to determine whether this principle applies in a similar way to data for females and males at other ages such as the 13 year threshold, and the 16 year threshold.

Ethical approval

In the UK ethical approval is not required for the analysis of previously published data which itself received ethical approval. It is usual under these circumstances to say “The authors confirm that the methods and procedures used in this project comply with UK law.”

Funding

The work was carried out as part of the intra-mural activity of employed and voluntary staff of the Department of Orthodontics, King's College London Dental Institute.

Conflict of interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

For the Excel spreadsheet comprising the Reference Data Set please go to: <http://dx.doi.org/10.1016/j.jflm.2014.05.004>. The Excel spreadsheet can also be obtained from the corresponding author.

References

1. Garn S, Lewis AB, Kerewsky RS. Genetic, nutritional, and maturational correlates of dental development. *J Dent Res* 1965;**44**(Suppl. 1):228–42.
2. Haavikko K. The formation and the alveolar and clinical eruption of permanent teeth. *Suom Hammaslaak Toim* 1970;**66**:103–70.
3. Demirjian A, Goldstein H, Tanner JM. A new system of Dental Age Assessment. *Hum Biol* 1973;**45**(2):211–27.
4. Bolanos MV, Manrique MC, Bolanos MJ, Briones MT. Approaches to chronological age assessment based on dental calcification. *Sci Int* 2000;**119**:97–106.
5. Liversidge HM, Lyons M, Hector MP. The accuracy of three methods of age estimation using radiographic measurements of developing teeth. *Sci Int* 2003;**131**:22–9.
6. Thevissen PW, Fieuws S, Willems G. Third molar development: measurements versus scores as age predictor. *Arch Oral Biol* 2011;**56**:1035–40.
7. Liversidge HM, Chaillet N, Mornstad H, Nystrom M, Rowlings K, Taylor J, et al. Timing of Demirjian's tooth formation stages. *Ann Hum Biol* 2006;**33**(4):454–70.
8. Thevissen PW, Fieuws S, Willems G. Human dental age estimation using third molar developmental stages: does a Bayesian approach outperform regression models to discriminate between juveniles and adults? *Int J Legal Med* 2010;**124**:35–42.
9. Roberts GJ, Parekh S, Petrie A, Lucas VS. Dental Age Assessment: a simple method for children and emerging adults. *Br Dent J* 2008;**204**(3):E7.
10. Fleiss JL. The statistical basis of meta-analysis. *Statist Meth Med Res* 1993;**2**:121–45.
11. Yadava M, Lucas VS, Roberts GJ. Dental Age Assessment: reference data for children at the 10-year-threshold. *Int J Legal Med* 2011;**125**:651–7 [E publication September 2010].
12. Chudasama PN, Roberts GJ, Lucas VS. Dental Age Assessment (DAA): a study of a Caucasian population at the 13 year threshold. *J Leg Med* 2012;**19**:22–8 [6].
13. Mitchell JC, Roberts GJ, Donaldson ANA, Lucas VS. Dental Age Assessment (DAA) for British caucasians at the 16 year threshold. *Sci Int* 2009;**189**:19–23 [4].
14. STATA. *Statistical software release 10*. Texas, USA: Stata Press; 2007, ISBN 1-59718-020-3.
15. Sterne JAC. *Meta-analysis in stata: an updated collection from the stata journal*. Texas: Stata Press; 2009, ISBN 978-1-59718-049-8.
16. Jelen B. *Microsoft excel*. USA: QUE Publishing; 2013. 978-0-789-74308-6.
17. Roberts GJ, Petrie A. *Dental Age Assessment: a practical approach*. Chapter 11 in Digital Forensic Science. IGI Global USA; 25 March 2011. 978-1-60960-483-7.
18. Petrie A, Sabin C. *Medical statistics at a glance*. 3rd ed. Oxford: Wiley-Blackwell; 2009, ISBN 978-1-4051-8051-1.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet* 1986; February 8th:307–10.
20. Everitt B. *Medical statistics from A to Z*. 2nd ed. Cambridge: Cambridge University Press; 2006, ISBN 0-521-68718-7.
21. Liversidge HM. Permanent tooth formation as a method of estimating age. Clinical aspects of dental morphology. *Front Oral Biol* 2009;**13**:153–7.